DCGAN-Based Data Augmentation for Tomato Leaf Disease Identification

Year: 2020 | Citations: 235 | Authors: Qiufeng Wu, Yiping Chen, Jun Meng

Abstract

Tomato leaf disease seriously affects the yield of tomato. It is extremely vital for agricultural economy to identify agricultural diseases. The traditional data augmentation methods, such as rotation, flip and translation, are severely limited, which cannot achieve good generalization results. To improve the recognition accuracy of tomato leaf diseases, a new method of data augmentation by generative adversarial networks (GANs) is proposed for leaf disease recognition in this work. Generated images augmented by deep convolutional generative adversarial networks (DCGAN) and original images as the input of GoogLeNet, this model can achieve a top-1 average identification accuracy of 94.33%. By adjusting the hyper-parameters, modifying the architecture of the convolutional neural networks, and selecting different generative adversarial networks, an improved model for training and testing 5 classes of tomato leaf images was obtained. Meanwhile, images generated by DCGAN not only enlarge the size of the data set, but also have the characteristics of diversity, which makes the model have a good generalization effect. We have also visually confirmed that the images generated by DCGAN have much better quality and are more convincing through the t-Distributed Stochastic Neighbor Embedding (t-SNE) and Visual Turing Test. Experiments with tomato leaf disease identification show that DCGAN can generate data that approximate to real images, which can be used to (1) provide a larger data set for the training of large neural networks, and improve the performance of the recognition model through highly discriminating image generation technology; (2) reduce the cost of data collection; (3) enhance the diversity of data and the generalization ability of the recognition models.